Translation and Quality
Editorial

Attila Görög
TAUS

ABSTRACT

The world is producing more and more content and needs increasing volumes to be translated. The only way to offer large amounts of information and goods in multiple languages fast while staying within reasonable budgets is by making a compromise and provide content with different levels of quality using new translation channels and translation technology. This differentiation will, in turn, require new ways of evaluating the quality of translated content. This issue has been co-edited by Attila Görög (TAUS) and Pilar Sánchez-Gijón (Tradumàtica Research Group), and focuses on translation quality, that remains a concept of major importance both in the translation industry and in academia.

Keywords: Translation quality, translation quality evaluation, translation quality estimation, translation industry, post-editing

Editorial

We live in the age of mass communication and personalized content: a plethora of information is freely available 24/7 tailored to our needs. We can find, read, buy, watch anything we want on the World Wide Web. Only the sky is the limit… and language. As for now, most information is still transmitted through language. 56% of web content is English only while there are billions of web pages, hundreds of thousands being added daily. This said, only a fraction of the world's population speaks English or wants to buy in English. If we want to ensure that online information is accessible for everyone, a multilingual Internet is essential as a mirror of reality, since multilingualism is in itself an intangible value, and also an asset to maintain as cultural and human heritage. What's more, there is also a business necessity.

A study published by the Harvard Business Review[1] found that 72 percent of all consumers spend most of their time on websites in their own languages. The same number is more likely to buy if product information is in their own language.[2] According to this same review more than half of the consumers are willing to pay more if you are willing to give them information in their own languages. A survey by Common Sense Advisory also supports this claim[3] and today it's becoming a common place to quote the former West German chancellor Willy Brandt who once said "If I am selling to you, I speak your language. If I am buying, dann

[1] Kelly, N. (on line). "Speak to Global Customers in their Own Language". In: *Harvard Business Review.* July-August 2012. <https://hbr.org/2012/08/speak-to-global-customers-in-t/> [Last accessed: 23/12/2014]
[2] Hall, S. H. (on line). "Popular Languages for Business Translation in 2015". In: *Smartling Blog.* 15th December, 2014. <http://www.smartling.com/2014/12/15/popular-languages-for-business-translation-in-2015/> [Last accessed: 23/12/2014]
[3] DePalma, D.A.; Hedge, V.; Stewart, R.G. (on line). "Can't Read, Won't Buy. How Translation Affects the Web Customer Experience and E-CommerceGrowth". In: *Common Sense Advisory Portal.* February 2014 <http://www.commonsenseadvisory.com/Portals/_default/Knowledgebase/ArticleImages/140214_CRWB_Abstract.pdf> [Last accessed: 23/12/2014]

müssen sie Deutsch sprechen".[4] The world is producing more and more content and needs increasing volumes to be translated. The only way to offer large amounts of information and goods in multiple languages fast while staying within reasonable budgets is by making a compromise and provide content with different levels of quality using new translation channels and translation technology. This differentiation will, in turn, require new ways of evaluating the quality of translated content.

This present issue of Tradumàtica starts with a quest for new approaches to and a new definition of translation quality. A group of authors from different professional backgrounds and affiliations sets out to explore "the Landscape of Translation" by trying to find workable definitions for translation, quality and translation quality in a sequence of three articles. This landscape is delimited by two axes—namely, scope and specifications—and provides a framework for discussing such questions as (a) the relationship between translation and localization and (b) the role of translation specifications. Within that landscape, the first article (**Melby et al.**) presents two opposing views about the definition of translation. The second article by **Fields et al.**, in turn, seeks to define quality describing five approaches to quality based on works from the discipline of quality management. Finally, in the third article, **Koby et al.** build on the first two and present two contrasting definitions of translation quality, a narrow one and a broad one. Throughout the three articles, the authors strongly disagree about which definition serves best the translation industry. Rather than pretending to be the end of the discussion, the authors of the trilogy hope to encourage a continuation of the debate that has taken place among them while in the process of writing the articles.

After these introductory thoughts on translation and quality, **Muzii** sets the stage in his article (cleverly entitled *The Red-pen syndrome*) by questioning the all-encompassing validity of the error-catching-and-assessment approach that is still one of the most widely used methods for assuring quality in the industry. One that is justified in some scenarios but almost impossible in others due to today's complex translation workflows and associated risks, tight budgets and increasing speeds of delivery. Still, most translation related standards in existence today follow the error-catching model. The article is a review of current translation quality standards and a proposal for student training from a historical perspective. Zooming in on the topic of student training, **Garcia** (*Training Quality Evaluators*) describes two opposing views on translation quality: the "enduring and critical" at one extreme, and the "ephemeral and inconsequential" at the other. The author proposes that, as content creation is becoming quicker, more agile and tailored to different user needs, translation training should also focus more on topics of efficiency, productivity and scalability of translation. This "fast and fit imperative" is also what teachers are facing today in the translation classroom when evaluating students. There is therefore a drive towards new approaches to translation quality identifying both translation quality *and* purpose in a more transparent and consistent way. Remaining in the realms of higher education, **Sanchez-Gijón**, examines different research directions concerning translation quality, and focuses on methodological issues that should be taken into account in order to contribute to the comparability of research outputs.

The third main topic of the present issue is concerned with tools and metrics for quality evaluation. Two main QE metrics are introduced and illustrated: the TAUS Dynamic Quality Framework (DQF) by **Görög** and the Multidimensional Quality Metrics (MQM) by **Lommel**. Both frameworks approach quality in a flexible way supporting the view that today's translation quality requirements change depending on content type, purpose and audience. Both frameworks are based on the assumption that traditional one-size-fits-all approaches do not satisfy buyers and vendors of translation services anymore. As a result, translation QE needs to embrace new methods that are differentiated based on customer needs. **Babych** approaches the topic of metrics from a different perspective. He addresses classes of methods for automatic evaluation of Machine Translation (MT) output quality. The main objective of methods and tools for automated evaluation of MT is to compute numerical

---

[4] The second half of the phrase translates: "then you must speak German." Brandt made an excellent point. In order to make your clients comfortable and instill trust, it's best to speak their language.

scores, which characterize the quality, or level of performance of specific MT systems. Automated MT evaluation scores are expected to agree (or correlate) with human intuitive judgments about certain aspects of translation quality, or with certain characteristics of usage scenarios for translated texts. The author describes the limitations of the existing metrics but also points out their value for professional translators and MT developers.

In the last part of this collection of articles, three authors zoom in on the practical side of translation quality. In the first article, **Aranberri** discusses the relationship between post-editing, productivity and quality. In recent years, progress in the field of MT has made large-scale adoption of MT usage possible. Not only because the quality of the leading systems has improved considerably, but also because the industry has recognized that even imperfect MT engines offer added value when facing current market demands. The author suggests that, consequently, the concept of quality is also changing. Standards and evaluation models created in the past for safeguarding the highest level of quality are losing popularity in favor of more flexible standards that take into account a variety of user needs. While human translation remains mandatory for certain fields such as medicine or the legal domain, MT combined with post-editing is becoming the norm when trying to meet the ever-growing demand for translations in the 21st century. But how to scale back translation size for quality evaluation (QE)? The answer can be found in the following paper where we read about *sampling*, a method increasingly used in the industry to obtain a smaller set of segments for quality review. One of the challenges when using sampling is to design and select samples carefully, so that they are representative of the content. Analyzing well-selected samples, for instance, can provide information on how a particular MT engine is performing and how to further improve the quality of the output. **De la Fuente** points out that randomly selected samples are usually less informative and he suggests that one should resort to systematic sampling, i.e. sampling performed using knowledge of the content. In his article, he describes how that knowledge can be acquired and applied in sample design. He also shares a list of recommendations for sample preparation based on his own experience. Sampling is also one of the topics of the use case article written by **Stoeller** who has helped develop VMware's review environment in which sampling also plays an important role. Stoeller describes the approach to quality measurement of human translations and the measurement of translator performance at VMware, a virtualization software company. The article focuses on the creation of the review environment using the TAUS content profiling tool and the DQF error typology model combined with sampling and performance trends.

This current issue is once again proof of the fact that in translation QE there is no black or white. There is no one-size-fits-all solution. Each content type, purpose or communication channel will require a different approach or combination of approaches. Moreover, available budgets, resources and customer requirements will determine the selection of the evaluation model(s) to be applied.

As we have mentioned earlier, translation quality remains of major importance both in the translation industry and in academia. Still, it is striking that these two spheres fail to collaborate on quality. There is, for instance, a general desire to standardize and compare evaluation results consistently. There is a need for defining and benchmarking quality levels. Besides, due to different constraints (time, money, human resources), QE does not always find its way into the translation workflow and into academic research. Industry and academia should work hand in hand to solve problems that currently hinder development in the area. For academic research to be successful in this field, large data sets are necessary. Evaluation data provides insights on quality expectations of different users and constitutes business intelligence for both buyers and vendors of translation and localization services.

Work in the area of quality evaluation has also been restricted by the lack of availability of relevant data to train metrics (such as post-edited MT output). This type of data is often abundant among providers and buyers of automatic translation, since they routinely need to assess translations for quality assurance. Research on better, reference-free automatic evaluation metrics would therefore greatly benefit from a closer relationship between industry

and academia. Feedback to researchers on the quality of the metrics and how they need to be further adapted to particular scenarios could also result in further improvements of such metrics. The benefits for the industry include better automatic metrics to support or minimize the need for human assessment and, potentially, better MT systems. Academia needs to obtain more feedback, information and requirements from the industry to better focus research activities on solutions to the problems that the industry is actually facing.

Frameworks and tools such as TAUS Dynamic Quality Framework (DQF) or the Multidimensional Quality Metrics (MQM) can facilitate such collaboration between industry and academia by providing systematic ways of collecting and storing quality assessment data (according to specific requirements for a given content type, audience, purpose, etc). Such data can then be directly used to improve existing evaluation algorithms. Additionally, automatic evaluation and quality estimation metrics could be integrated into such platforms to support human evaluation.

This recent issue of Tradumàtica is a good example of collaboration between industry and academia. The two editors as well as the numerous authors of this issue are from different backgrounds. The topics selected are of interest for a wide variety of readers. The contributions offer both theoretical and practical approaches to the subject of translation quality opening the way for collaboration between the two camps. Finally, we hope that this present issue will serve as a first step in facilitating and moving forward the dialogue on translation quality to benefit both translation research and practice. Enjoy reading!